DIGITALIZATION OF THE EU COUNTRIES: A CLUSTERWISE ANALYSIS

Jani KINNUNEN [a]***, Armenia ANDRONICEANU*** [b]***, Irina GEORGESCU*** [b*]

[a] *Abå Akademi, Turku, Finland*
[b] *The Bucharest University of Economics Studies, Bucharest, Romania*

ABSTRACT
European Union is heading towards a digital single market based on its Digital Single Market strategy accepted in 2015. European countries have started their national digitalization strategies from very different levels and the gap is considerable between frontrunners and challengers. In this paper, multivariate data analysis techniques are used: 29 European countries are firstly grouped by clustering analysis using ten variables of different aspects of digitalization. Five clusters of similar countries are obtained. The clusters, as well as, the countries within each cluster show consistent performance roughly over all indicators. Secondly, principal component formulas are constructed based on principal component analysis. This is done for the EU as a whole and, separately, for each cluster.

KEYWORDS: *cluster analysis, digitalization, European Union, principal component analysis*

1. INTRODUCTION

The world is at the threshold of the fourth industrial revolution, which has already begun. This fact changes all enterprises, regardless of their sector of activity (Shpak et al., 2019) and even countries. Digitalization is one of the most important features of this revolution. It has multiple aspects. The actors on the field are various ranging from transnational bodies, such as the EU, OECD and World Bank to numerous groups of lobbyist and activists, from governments and enterprises to individuals in a society and workers in companies. Also, the drivers of digitalization are diverse, such as the use of the Internet, the state of digital skills of individuals and workers, or digital services used by citizens to improve their quality of life (Ragnedda, M. and H. Kreitem, 2018; Belas et al., 2019; Kral et al., 2018; Mura et al., 2018; Stverkova et al., 2018; Ślusarczyk, 2018; Ključnikov et al., 2019). One should also add that implementation of Industry 4.0 requires continuous innovation and education (Durana et al., 2019) which cannot be done without digitalization. Digitalization is expected to change economic and social life. The digitalization of business and administration processes comes with benefits for governments, companies, and individuals (Balcerzak & Pietrzak, 2017; Sebestova et al., 2018; Ungerman et al., 2018; Tvaronavičienė, 2019). New jobs will be created and many jobs will be transformed. The economic benefits have been estimated at €415 billion a year, at the same time with new job creation for hundreds of thousands (European Commission, 2017).
In 2015, the European Commission approved the Digital Single Market strategy (European Commission, 2015) to create a European digital society. Its aim is to ensure that Europe's economy, industry and society are aligned with the new digital era. 20 out of 29 EU countries have adopted digital strategies or set digital agendas, for example, Austria, Germany and Slovenia. Androniceanu et al. (2017) analyze the opinions of young people on the impact of socio-economic digitalization on consumers' perception and behavior. The results of the study show that young people are receptive

* Corresponding author. E-mail address: irina.georgescu@csie.ase.ro

to the services of the digital society, in spite of the fact that Romania places last on the 28th place according to the economy and society index DESI. In another study, Burlacioiu et al. (2018) apply the Ward hierarchical clustering to divide the 28 EU member states into two clusters. According to their study, the first cluster contains Denmark, Sweden, Finland, Estonia, Luxembourg, the Netherlands and the United Kingdom, while Romania places in the second cluster. Further, the authors apply principal component analysis, PCA, and remark that Romania and Bulgaria have the lowest indicators of digital indicators on both PCA axes.

Disparities between countries and clusters of countries are considerably large within European Union (Mura et al., 2017; Lazaroiu et al.2018). In this paper we propose a clustering approach with respect to digitization indicators in order to find disparities between countries.

This paper is organized as follows. In section 2, the data and descriptive statistics for the 29 European countries (the 28 member countries and Norway as part of the Northern European highly digitalized countries) are presented. Section 3 covers the two-step multivariate approach of this study. Section 3 starts by taking a closer look of clusters, seen also in Figure 1, obtained by clustering analysis, by first reporting the cluster wise indicator performance of all clusters. Then, after explaining certain aspects of the other used method, PCA is conducted and the results for the five clusters of countries are reported.

Section 3 will conclude the research and suggest future research opportunities. Some conclusions on the clusters are presented below. We remark (see Figure 1) the high-performing countries such as Denmark, Finland, Norway, Sweden, Netherlands, from cluster 1 in Figure 1, Luxembourg, Estonia and UK from cluster 2, or Belgium and Ireland from cluster 3; and some countries are in the bottom by most or all indicators, specifically, Romania and Bulgaria (cluster 5 in figure 1), or Croatia, Cyprus, Greece, Hungary, Italy, Latvia, Lithuania, Poland and Portugal (cluster 4 in Figure 1). More efforts are therefore needed from less developed countries to achieve convergence in the digital single market.

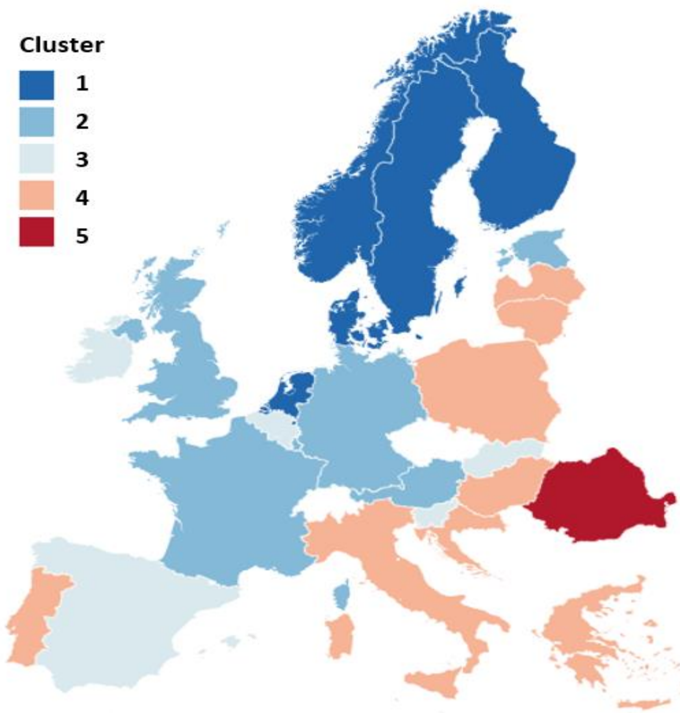


Figure 1. Country clusters from the most digitalized (blue) to the least digitalized countries (red)

Source: own computations

There are both opportunities and threats seen in the digitalization process. There are barriers for development, e.g. individuals still have limited access even to the Internet in some regions, and customers can have limited access to goods and services online. Entrepreneurs and start-ups may not see great opportunities to invest in digitalization in all areas. Businesses and also governments have obstacles to benefit from digitalization. The Digital Single Market strategy of the EU aims to lower these barriers and boost moving from national markets to one larger single digital European market. A competitive environment of a single market can generate transparency, accessibility, cheaper products and services, and help create innovative new solutions and facilitate the free movement of goods and services, whether digital or physical ones.

Most European countries are involved in e-government activities. They transfer online parts of administrative tax procedures, electronic tax audits and data gathering from taxpayers (Shkarlet et al.2019). Austria, Lithuania, Slovenia, Hungary, Portugal and Poland have implemented e-audit and e-reporting including e-invoicing. Luxembourg, France, Germany, Greece, Netherlands have implemented e-audit. Croatia, Italy, the Czech Republic, Spain and Estonia use e-reporting. Slow progress in these areas are seen in Romania, Bulgaria and Malta.

Even that the same countries, specifically, Romania and Bulgaria place in a lower position compared to other European countries by all digitalization indicators, these countries also offer the greatest potential. Recent McKinsey report (Novak et al., 2018) discusses how digitization can become the next growth engine for Romania and Bulgaria together with Croatia, Czech Republic, Latvia, Lithuania, Poland, Slovakia, Slovenia and Hungary. These "Digital Challengers" demonstrate a strong growth potential in the digital economy and their rival "Digital Frontrunners" such as Belgium, Denmark, Estonia, Finland, Ireland, Luxembourg, the Netherlands, Norway and Sweden.

Methodologically, as Vezzani et al. (2018) from European Commission's research center point out, the political and economic challenges faced by the EU call for evidence-based policies. This means developed analyses, which are not only based not on country-level data, but "rather on regional and sub-regional knowledge". This is also a motivation for this paper's clustering analysis. Further, Vezzani et al. (2018) criticize the use of national averages in analyses as they may give a misleading picture of countries with large disparities between different regions and areas; national averages can also prevent the identification of emerging trends in certain socio-economic indicators (Guilherme Nunes et al., 2018). Only a detailed analysis of data at regional and local level can bring these insights.

2. DATA

The data is collected for the 29 European countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and United Kingdom.

The used ten indicators are extracted from Eurostat's *Digital Economy and Society database* for year 2018. All available digitalization indicators in the database were selected. The ten indicators can be divided into three types: (a) *digitalization of individuals*: (a1) who use internet; (a2). who make purchases online and (a3) who use government services, (b) *digitalization of workers* (b4) receiving trainings and (b5) using ICT in enterprises, and (c) *digitalization of enterprises*, (c6) which have websites; (c7) which sell products and services online, (c8) which use big data; (c9) which use cloud computing, and (c10) which recruit ICT specialists.

The descriptive statistics for the indicators included in the analysis show that their standard deviations do not differ very much (Table 1) and the variations within clusters will be seen even significantly lower. The more the standard deviations differ from one another, the more the countries differ with respect to that indicator.

Table 1. Indicators and descriptive statistics

	N	Minimum	Maximum	Sum	Mean		Std. Deviation	Variance
Internet use	29	67	98	2477	85.41	1.508	8.122	65.966
Internet purchases	29	13	77	1285	44.31	3.291	17.722	314.079
E-government activities	29	9	92	1626	56.07	3.752	20.208	408.352
ICTskills	29	2	27	265	9.14	1.278	6.88	47.337
ICT impact	29	2	19	242	8.34	0.942	5.073	25.734
Corporate website	29	44	96	2219	76.52	2.315	12.469	155.473
E-commerce	29	4	35	490	16.9	1.465	7.889	62.239
Big data analysis	29	5	24	361	12.45	0.959	5.166	26.685
ICT Recruitment	29	4	13	256	8.83	0.479	2.578	6.648
Cloud computing	29	8	65	863	29.76	2.807	15.118	228.547

Source: own computations

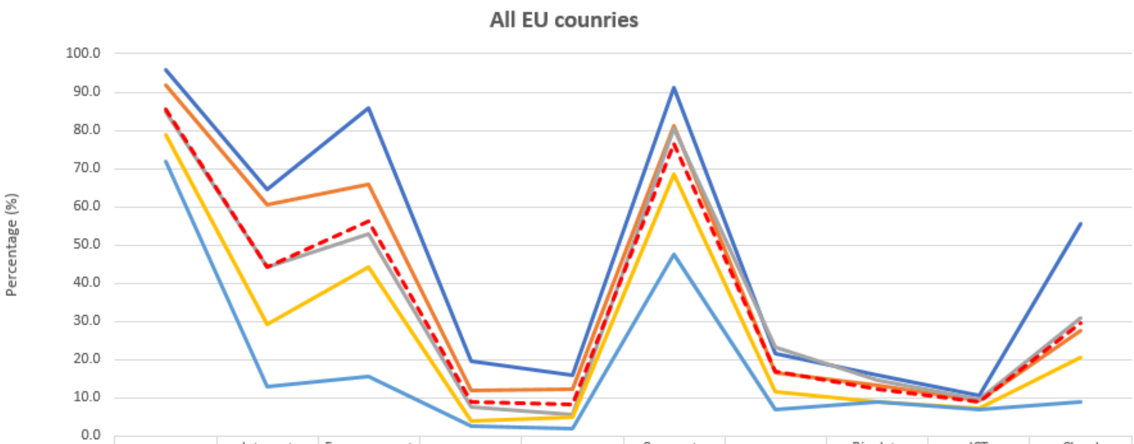
3. CLUSTERING AND PRINCIPAL COMPONENT ANALYSIS

The first step in the analysis is to conduct k-means clustering, one of the most popular clustering methods. k centroids are established in order to define k clusters. An object belongs to a cluster if it is closer to that cluster's centroid than to any other centroid. The algorithm finds the best clustering by alternating two steps:
Step 1: The objects are assigned to clusters in accordance with the current centroids;
Step 2: The new centroids are chosen based on the current assignment of objects to clusters.
We decided that for this analysis, the most appropriate number of clusters is 5. The 29 EU countries are grouped into following five clusters (cf. Figure 1):

Cluster 1 (5 countries): Denmark, Finland, Netherlands, Norway, and Sweden;
Cluster 2 (6 countries): Austria, Estonia, France, Germany, Luxembourg, and UK;
Cluster 3 (7 countries): Belgium, Czech Republic, Ireland, Malta, Slovakia, and Slovenia;
Cluster 4 (9 countries): Croatia, Cyprus, Greece, Hungary, Italy, Latvia, Lithuania, Poland and Portugal; and
Cluster 5 (2 countries): Bulgaria and Romania.

This paper continues the research by Androniceanu et al. (2019) with new findings on cluster analysis using the same data set. The mean values of indicators for each cluster (CL1-CL5) and all countries are shown in Figure 2. We remark that cluster 5 (CL5) contains only Bulgaria and Romania (see also Figure 7 in the end of the section), and cluster 1 (CL1) contains mainly the Nordic countries, Denmark, Finland, Norway, and Sweden with Netherlands. From Figure 2, one sees that CL5 distinguishes itself by the poorest mean of all digitalization indicators. At the opposite pole, CL1 has the highest value of all digitalization indicators except the (7.) e-commerce share of sales (21.2%) is slightly higher on average for CL3 (23.3%) in which Czech Republic, Belgium and Ireland show shares as high as 29-35%.
CL3 is seen to be very close to the average of all countries by every indicator; only by e-commerce (7.) the all-country average is significantly different (16.9%). CL4 is the second lowest, only above CL5, by every indicator. Only CL4 doesn't seem to hold its (second highest) position consistently through all the indicator means. It is between CL1 and CL3 by all indicators concerning individuals in general and workers at enterprises (1.-5.) and also by corporate websites (6.), but below both CL1 and CL3 by enterprise-related indicators (7.-10.) of e-commerce, use of big data and cloud computing, as well as, by recruiting ICT specialists.

In general, the clusters perform consistently in comparison to other clusters over all indicators under analysis and the cluster averages are not very far from other clusters' means except by e-government activities (3.), internet purchases (2.) and corporate use of cloud computing (10.); the three variables also show the biggest standard deviations in Figure 1, 20.2, 17.7, and 15.1, respectively. E-government activity is measured by individuals' use of electronic government services, which obviously depends heavily on government efforts to offer e-government services. E-government activities are measured by 81.5% of individuals, while only 15.5% in CL5, the average of all 29 countries being 56.1%; 66.5% of CL1 individuals did internet purchases, while in CL3 only 13% (total average being at 44.3); As many as 53.2% of CL1 enterprises used cloud computing, while only 9% in CL5 (total average is 29.8%).
Further, we can note, that we cannot see very notable differences between counties within clusters in most of the variables, but again significant differences are seen in e-government activities (3.), specifically in CL1, CL4, and somewhat in other clusters, and the corporate use of cloud computing (10.), as well as, the internet purchases by individuals (2.) in all clusters except CL5 (Romania and Bulgaria).



	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
CL1-avg	95.8	64.4	86.0	19.6	15.8	91.2	21.6	16.0	10.8	55.4
CL2-avg	91.7	60.7	65.8	11.8	12.3	81.2	16.5	13.2	9.2	27.5
CL3-avg	84.9	44.3	53.0	7.7	5.7	80.6	23.3	14.6	9.6	31.0
CL4-avg	78.9	29.2	44.3	4.1	5.0	68.6	11.8	9.1	7.3	20.7
CL5-avg	72.0	13.0	15.5	2.5	2.0	47.5	7.0	9.0	7.0	9.0
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 2. Mean value of indicators – All EU countries
Source: own computations

The second step in this study's applied approach is to run principal component analysis, PCA (cf. Jolliffe, 2002). This is done using Tanagra software (Rakotomalala, 2013). The correlation matrix (not reported here) shows that the indicators are strongly interrelated, meaning informational redundancies, which supports the use of PCA. There are a set of tests that are used with PCA. *Kaiser's Measure of Sampling Adequacy (MSA)*measures the sampling adequacy, which varies between 0 and 1. The values closer to 1 are better and the value of 0.6 is the suggested minimum. For our dataset, MSAs are greater than 0.6. This is acceptable for applying PCA.
Tanagra software performs PCA based on the correlation matrix. For the EU as a whole, according to the principle of proportion of explained variance, we will retain the first two principal components (PCs), which explain together 75.26% of the total variability as seen in Table 2. There are also other tests to help choose the number of PCs. For example, the same fact is confirmed by *Kaiser-Guttman principle*, by which if the variables are independent, the eigenvalues for all the factors are1. Thus, we select the first two PCs, corresponding to the first two eigen values greater than 1. The significance of the retained PCs is given also by *Karlis-Saporta-Spinaki test* (Karlis et al., 2003). This rule is more

restrictive than Kaiser-Guttman rule; it selects the PCs with the eigenvalues λ greater than 1, which verify

$$\lambda > 1 + 2\sqrt{\frac{p-1}{n-1}},$$

where p denotes the number of variables and n denotes the number of objects. For this dataset of 10 variables and 29 countries dataset, we have λ greater than 1:

$$\lambda > 1 + 2\sqrt{\frac{p-1}{n-1}} = 1 + 2\sqrt{\frac{10-1}{29-1}} = 2.13.$$

Table 2 shows above the typical factor loadings for PCs. Factor loadings or communality estimates represent the proportions of each variable's variance which can be explained by the factors. The colored loading factors indicate that the factors are dominated by the corresponding variables. Top of Table 2 has loading approximations larger than 0.5 for all factors; that's why we have further run PCA using *varimax method* to get more accurate results. Both methods show the first two PCs to explain exactly the same of the total variance although the latter shift part of the emphasis from PC1 (explains 48.7% of total variation of all variables) to PC2 (explains 26.5%).

Table 2. Factor coefficients (loadings) for principal components – All EU countries

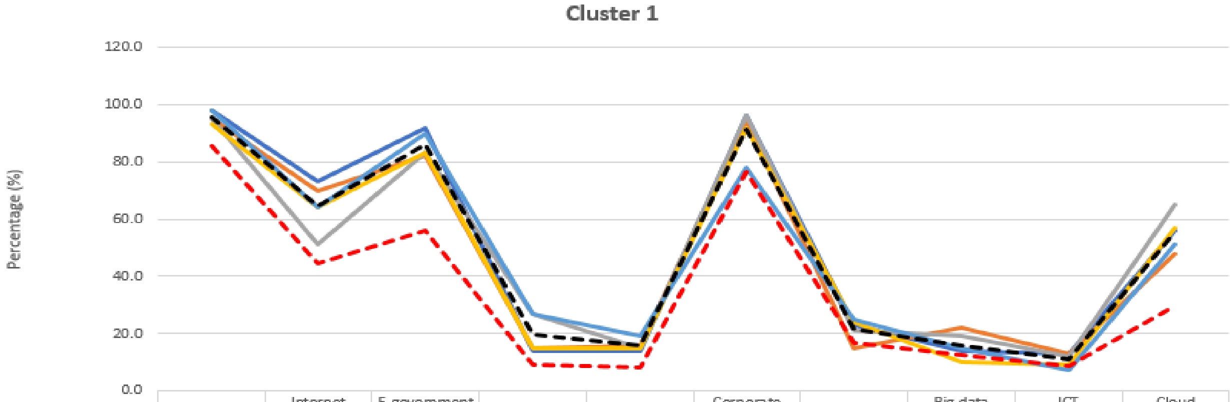
All EU countries		Internet-purchase	Internet-use	Website	ICT-impact	Cloud-computing	E-government	ICT-skills	Recruitment	Big-data-analysis	E-commerce	Var. Expl.
PC1	Corr.	0.911	0.907	0.871	0.870	0.866	0.864	0.846	0.626	0.614	0.576	*65 % (65 %)*
PC2	Corr.	-0.119	-0.226	0.010	-0.348	0.120	-0.261	-0.185	0.571	0.564	0.313	*10 % (75 %)*
PC3	Corr.	0.003	-0.015	-0.049	0.130	-0.050	-0.068	0.171	0.201	0.311	-0.727	*7 % (82 %)*
PC1-varimax	Corr.	0.884	0.830	0.867	0.811	0.919	0.726	0.313	0.209	0.216	0.662	*48 %(65 %)*
PC2-varimax	Corr.	0.303	0.396	0.250	0.305	0.181	0.482	0.576	0.807	0.820	0.571	*27% (75 %)*

Source: own computations

Using the factor coefficients in the bottom of Table 2, we formulate PCs:

*PC1 = 0.919*ICT-impact + 0.884*Internet-use + 0.867*E-government + 0.83*Internet-purchases + 0.811*ICT-skills + 0.726*Corporate-website + 0.662*Cloud-computing*
*PC2 = 0.807*Big-data + 0.82*ICT-recruitment + 0.576*E-commerce*

PC1 and PC2 cannot have same indicators as they must be independent of each other. Thus, not surprisingly, the three indicators of PC2 showed the lowest correlations with all the other variables in the chart of Figure 1. PC2 is formed by the three enterprise-variables (7.-9.) of using big data, recruiting ICT specialists, and doing e-commerce as a share of sales. The coefficients are very high (between 0.82- 0.919) for 8 out of our 10 variables, and moderate 0.662 for (10.) cloud computing, and the lowest, 0.5569 for (7.) e-commerce. The highest coefficient of 0.919 is seen for (5.) ICT impact, i.e. the use of computers, software or applications in enterprises explain most of the total variance.



	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
Denmark	98.0	73.0	92.0	14.0	14.0	96.0	23.0	14.0	13.0	56.0
Netherlands	95.0	70.0	82.0	15.0	16.0	94.0	15.0	22.0	13.0	48.0
Finland	95.0	51.0	83.0	27.0	15.0	96.0	21.0	19.0	12.0	65.0
Sweden	93.0	64.0	83.0	15.0	15.0	92.0	24.0	10.0	9.0	57.0
Norway	98.0	64.0	90.0	27.0	19.0	78.0	25.0	15.0	7.0	51.0
CL1-avg	95.8	64.4	86.0	19.6	15.8	91.2	21.6	16.0	10.8	55.4
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 3. Mean value of indicators – Cluster 1

Source: own computations

Table 3. Factor coefficients (loadings) for principal components – Cluster 1

Cluster 1		Corporate-website	ICT-Recruitment	ICT-impact	E-commerce	E-government-activities	Internet-use	ICT-skills	Internet-purchases	Cloud-computing	Big-data-analysis	Var. Expl.
PC1	Corr.	0.939	0.867	-0.793	-0.723	-0.594	-0.542	-0.536	0.017	0.230	0.411	*39 % (39 %)*
PC2	Corr.	-0.034	0.255	-0.003	-0.220	0.492	0.482	-0.658	0.978	-0.770	0.030	*26 % (65 %)*
PC3	Corr.	0.292	-0.013	-0.595	0.655	0.397	-0.016	-0.357	0.129	0.478	-0.823	*21 % (86 %)*

Source: own computations

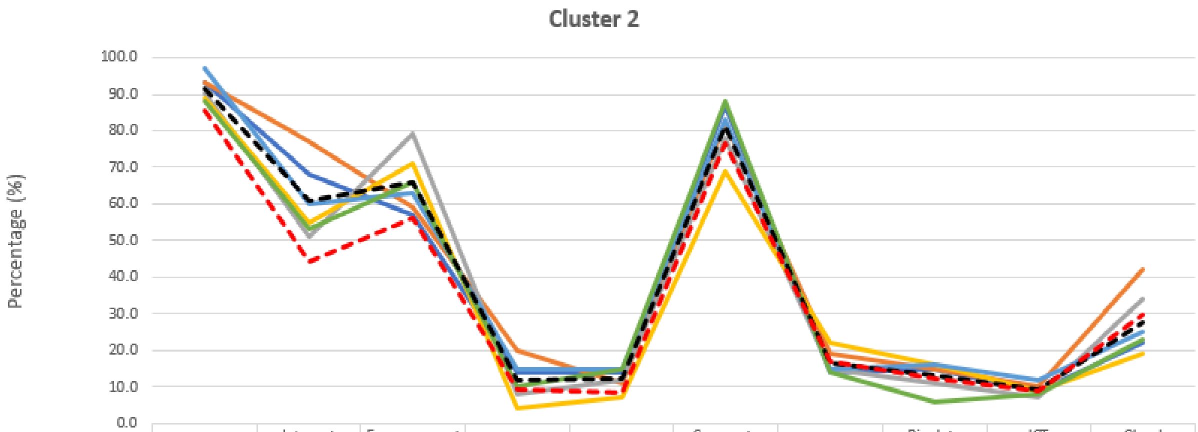
From Table 3, the principal components, PCs, can be constructed for cluster 1:

*PC1= 0.939*Corporate-website + 0.867*ICT-Recruitment - 0.792*ICT-impact - 0.722*E-commerce - 0.593*E-government - 0.542*Internet-use - 0.536*ICT-skills*
*PC2 = 0.978*Internet-purchases - 0.770*Cloud-computing*
*PC3= - 0.823*Big-data*

In Cluster 1, the most of the indicators dominate positively (Corporate-website, ICR-Recruitment) or negatively (ICT-impact, E-commerce, E-government, Internet-use, and ICT-skills) PC1. So, PC1 can be viewed as a measure of quality of corporate website and ICT recruitment. The highly negative coefficients in PC1 give a problematic interpretation of this PC, given that the Nordic countries are leaders in the development and exploitation of new technologies.

Table 4 shows the factor coefficients for the PCs of cluster 2:

*PC1= - 0.943*ICT-skills + 0.849*E-government - 0.821*Internet-use - 0.789*Internet-purchases - 0.703*ICT-Recruitment - 0.649*Corporate-website - 0.555*ICT-impact*
*PC2 = - 0.885*E-commerce - 0.839*Big-data*
*PC3 = - 0.850*Cloud-computing*



	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
Germany	93.0	68.0	57.0	14.0	14.0	87.0	14.0	15.0	9.0	22.0
UK	93.0	77.0	59.0	20.0	11.0	82.0	19.0	15.0	10.0	42.0
Estonia	90.0	51.0	79.0	8.0	12.0	78.0	15.0	11.0	7.0	34.0
France	89.0	55.0	71.0	4.0	7.0	69.0	22.0	16.0	9.0	19.0
Luxembourg	97.0	60.0	63.0	15.0	15.0	83.0	15.0	16.0	12.0	25.0
Austria	88.0	53.0	66.0	10.0	15.0	88.0	14.0	6.0	8.0	23.0
CL2-avg	91.7	60.7	65.8	11.8	12.3	81.2	16.5	13.2	9.2	27.5
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 4. Mean value of indicators – Cluster 2

Source: own computations

Table 4. Factor coefficients (loadings) for principal components – Cluster 2

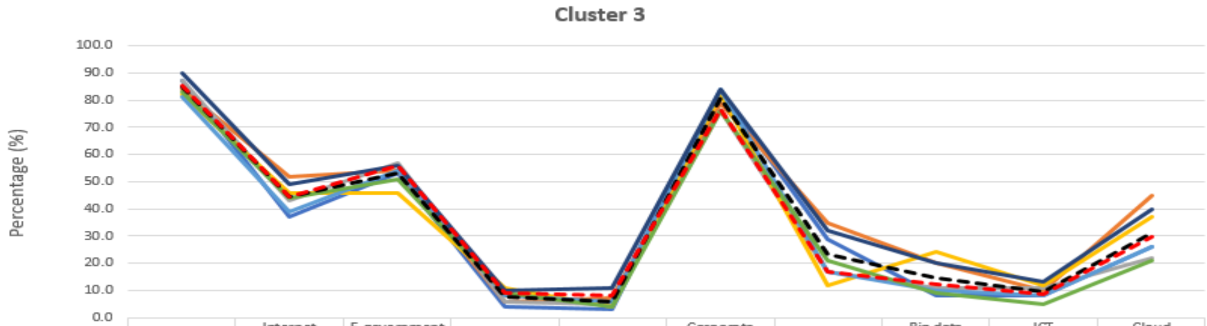
Cluster 2		ICT-skills	E-government-activities	Internet-use	Internet-purchases	ICT-Recruitment	Corporate-website	ICT-impact	E-commerce	Big-data-analysis	Cloud-computing	Var. Expl.
PC1	Corr.	-0.944	0.849	-0.821	-0.789	-0.703	-0.649	-0.555	0.333	-0.348	-0.326	*45 % (45 %)*
PC2	Corr.	-0.021	0.072	-0.238	-0.390	-0.422	0.738	0.794	-0.885	-0.839	-0.130	*31 % (75 %)*
PC3	Corr.	-0.322	-0.092	0.314	-0.360	0.443	-0.051	0.196	-0.138	0.273	-0.850	*14 % (89 %)*

Source: own computations

For cluster 2, PC1 is governed by negatively dominating variables. Only the e-government activity has a positive coefficient. Countries in cluster 2 are relatively high in that measure compared to the EU average, specifically Estonia, as well as, France. Cluster 2 is very high on internet purchases too, it is seen particularly in PC1. PC2 is positively dominated by corporate websites and ICT-related workers and negatively by online sales. Thus, individuals shop heavily online, but it is implied that not from this cluster's enterprises. PC3 is negatively dominated by only the use of cloud computing reflecting under EU-average performance (Figure 4) except by France and UK.

From Table 5, the coefficients are picked up for the PCs for cluster 3:

*PC1 = 0.906*Cloud-computing + 0.854*Big-data + 0.818*ICT-impact + 0.817*Internet-purchases*
*PC2 = - 0.891*E-government - 0.716*Internet-use + 0.709*ICT-skills - 0.672*E-commerce*
*PC3 = - 0.876*Corporate-website*



	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
Czechia	87.0	37.0	53.0	4.0	3.0	83.0	29.0	8.0	8.0	26.0
Ireland	84.0	52.0	54.0	6.0	7.0	79.0	35.0	20.0	10.0	45.0
Spain	87.0	43.0	57.0	6.0	5.0	76.0	17.0	11.0	11.0	22.0
Malta	82.0	46.0	46.0	11.0	4.0	82.0	12.0	24.0	12.0	37.0
Slovenia	81.0	39.0	54.0	8.0	6.0	84.0	17.0	10.0	8.0	26.0
Slovakia	83.0	44.0	51.0	9.0	4.0	76.0	21.0	9.0	5.0	21.0
Belgium	90.0	49.0	56.0	10.0	11.0	84.0	32.0	20.0	13.0	40.0
CL3-avg	84.9	44.3	53.0	7.7	5.7	80.6	23.3	14.6	9.6	31.0
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 5. Mean value of indicators – Cluster 3

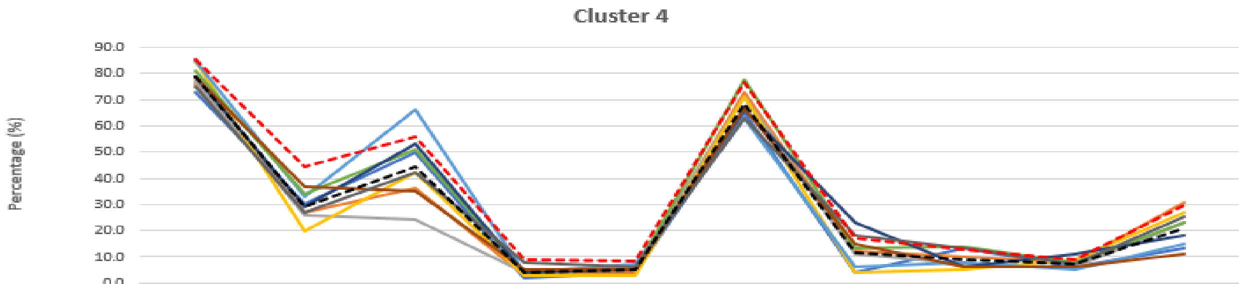
Source: own computations

Table 5. Factor coefficients (loadings) for principal components – Cluster 3

Cluster 3		Cloud-computing	Big-data-analysis	ICT-impact	Internet-purchases	ICT-Recruitmen	E-government-activities	Internet-use	ICT-skills	E-commerce	Corporate-website	Var. Expl.
PC1	Corr.	0.906	0.854	0.818	0.817	0.796	0.108	0.405	0.395	0.474	0.339	*42 % (42 %)*
PC2	Corr.	0.128	0.476	-0.308	0.119	0.047	-0.891	-0.716	0.709	-0.672	0.061	*26 % (68 %)*
PC3	Corr.	0.066	0.076	-0.106	0.559	-0.242	0.004	-0.130	-0.082	0.163	-0.876	*12 % (80 %)*

Source: own computations

In cluster 3, PC1 is dominated strongly positively by the corporate use of cloud computing and big data, and recruiting ICT specialists by workers' use and working on software and ICT (ICT-impact) together with individuals' internet use. PC1 can be described as an enterprise-driven component. PC2 instead is strongly negatively dominated by the indicators of E-government, Internet-use and E-commerce and PC2 increases with workers' ICT skills. PC3 is fully dominated negatively by the indicator of corporate websites. Again, large negative values for PC2 and PC3 cause problems for the interpretation.



	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
Greece	73.0	30.0	50.0	2.0	4.0	65.0	4.0	13.0	6.0	13.0
Croatia	76.0	27.0	36.0	3.0	4.0	73.0	12.0	10.0	8.0	31.0
Italy	77.0	26.0	24.0	4.0	6.0	71.0	11.0	7.0	6.0	23.0
Cyprus	85.0	20.0	42.0	3.0	3.0	71.0	4.0	5.0	9.0	27.0
Latvia	85.0	33.0	66.0	4.0	7.0	63.0	6.0	8.0	5.0	15.0
Lithuania	81.0	34.0	51.0	4.0	5.0	78.0	13.0	14.0	8.0	23.0
Hungary	79.0	29.0	53.0	4.0	5.0	66.0	23.0	6.0	11.0	18.0
Poland	79.0	37.0	35.0	5.0	5.0	67.0	15.0	6.0	6.0	11.0
Portugal	75.0	27.0	42.0	8.0	6.0	63.0	18.0	13.0	7.0	25.0
CL4-avg	78.9	29.2	44.3	4.1	5.0	68.6	11.8	9.1	7.3	20.7
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 6. Mean value of indicators – Cluster 4

Source: own computations

From Table 6, we get the PCs for cluster 4:

*PC1 = - 0.779*ICT-impact - 0.743*Internet-purchases + 0.740*Cloud-computing*
*+ 0.643*Corporate-website + 0.600*ICT-Recruitment*
*PC2 = 0.790*E-commerce + 0.716*ICT-skills - 0.566*Internet-use*
*PC3 = - 0.680*Big-data*
*PC4 = - 0.547*E-government*

Table 6. Factor coefficients (loadings) for principal components – Cluster 4

Cluster 4		ICT-impact	Internet-purchases	Cloud-computing	Corporate-website	ICT-Recruitmen	E-commerce	ICT-skills	Internet-use	Big-data-analysis	E-government-activities	Var. Expl.
PC1	Corr.	-0.779	-0.743	0.740	0.643	0.600	-0.169	-0.437	0.089	-0.209	-0.416	*29 % (29 %)*
PC2	Corr.	0.189	-0.072	0.381	0.019	0.269	0.790	0.716	-0.566	0.275	-0.464	*20 % (49 %)*
PC3	Corr.	0.160	-0.085	-0.044	-0.206	0.543	0.448	0.231	0.634	-0.680	0.290	*16 % (65 %)*
PC4	Corr.	0.150	-0.410	0.027	-0.386	-0.428	-0.258	0.179	-0.031	-0.512	-0.547	*12 % (77 %)*

Source: own computations

In cluster 4, PC1 increases with Cloud-computing, Corporate-website and ICR-recruitment, indicating that the countries of this cluster have high values of these indicators, and lower values of Internet Purchases and ICT-impact.
For cluster 5 of Romania and Bulgaria, we rely only on the indicator values as the cluster is not of adequate size for pursuing principal component analysis. Romania and Bulgaria are the weakest performers with EU by almost all digitalization indicators under study. Bulgaria shows the average EU level performance in ICT recruitment, higher than Romania. Also, Bulgaria shows more e-government activity than Romania; otherwise they show very similar digitalization stages, when compared to each other. As a policy, the Romanian government adopted the Digital Agenda for Romania - 2020. The areas of action of the National Strategy for the Digital Agenda are: e-Government, Interoperability, Cyber Security, Cloud Computing and Social Media, ICT in education, culture and health, ICT in e-commerce, and research, development and innovation, Broadband and digital infrastructure services. The application of these policies will lead to the increased use of Internet, the increased access to public services, E-commerce promotion, supporting research and innovation in ITC.

	Internet use	Internet purchases	E-government activities	ICT skills	ICT impact	Corporate website	E-commerce	Big data analysis	ICT Recruitment	Cloud computing
Bulgaria	67.0	13.0	22.0	2.0	2.0	51.0	5.0	7.0	10.0	8.0
Romania	77.0	13.0	9.0	3.0	2.0	44.0	9.0	11.0	4.0	10.0
CL5-avg	72.0	13.0	15.5	2.5	2.0	47.5	7.0	9.0	7.0	9.0
All-avg	85.4	44.3	56.1	9.1	8.3	76.5	16.9	12.4	8.8	29.8

Figure 7. Mean value of indicators – Cluster 5
Source: own computations

3. CONCLUSIONS

The results show the current (2018) state of digitalization in the 28 EU countries and Norway. The most digitalized countries of cluster 1 including the Nordic countries of Denmark, Finland, Norway and Sweden together with Netherlands perform better than others in all considered digitalization indicators, while Romania and Bulgaria (cluster 5) show poor performance. Cluster 4 countries (Croatia, Cyprus, Greece, Hungary, Italy, Latvia, Lithuania, Poland and Portugal) do not perform much better except on e-government activities. However, a lot of expectations is put on low-performing countries as they have the greatest potential to develop, while they also show some recent development. For example, McKinsey (Novak et al., 2018) estimates the potential economic benefits of digitization up to € 200 billion in additional GDP by 2025 for the Central and Eastern Europe (CEE).

The paper showed how cluster analysis finds consistently similarly performing countries in all the obtained five clusters, while principal component analysis showed the key variables in each cluster and revealed significant differences between clusters.

For future research opportunities, one can use discriminant analysis and composite index for a set of digitalization variables. A composite index approach would work well also with time-series data, which would bring dynamic perspective into analysis.

REFERENCES

Androniceanu, A., Georgescu, I. & Kinnunen, J. (2019). *Digitalization Clusters within the European Union*, IBIMA, 33rd IBIMA Conference, Granada, Spain 10-11 April, 2019, 1321-1331

Androniceanu, A., Burlacu S., Drăgulănescu, I. V., Nicolae, E. E. (2017). *New Trends of Business Digitalization in Romania and the Behaviour of Young Consumers*, BASIQ 2017: New Trends in Sustainable Business and Consumption, Graz- Austria, May-June 2017, 27-35

Balcerzak, A. P., Pietrzak, M. B. (2017). Digital economy in Visegrad countries. Multiple-criteria decision analysis at regional level in the years 2012 and 2015. *Journal of Competitiveness*, *9*(2), 5-18. DOI: 10.7441/joc.2017.02.01

Belas, J., Belas, L., Cepel, M., Rozsa, Z. (2019). The impact of the public sector on the quality of the business environment in the SME segment, *Administratie si Management Public*, (32),18-31. DOI: 10.24818/amp/2019.32-02

Burlacioiu, C., Moise, I., Boboc, C., Croitoru, E. O. (2018). *Digital Technology Trend in Romania and Its Impact on Young Segment*, Proceedings of the 12th International Management Conference "Management Perspectives in the Digital Era", November 2018, Bucharest, Romania, 824-835

Durana, P. Kral, P., Stehel, V., Lazaroiu, G., Sroka, W. (2019). Quality culture of manufacturing enterprises: a possible way to adaptation to industry 4.0. *Social Sciences*, 8, 124. doi:10.3390/socsci8040124

European Commission (2015). *A Digital Single Market Strategy for Europe – Analysis and Evidence*. Commission Staff Working Document SWD (2015) 100. European Union.

European Commission (2015). *A Digital Single Market Strategy for Europe*. A Communication PaperSWD(2015) 192. European Union.

European Commission (2017). *European Commission Concept Paper on Digitisation, Employability and Inclusiveness*. European Union.

Guilherme Nunes, F., Martins, L. M., Mozzicafreddo, J. (2018). The influence of service climate, identity strength, and contextual ambidexterity upon the performance of public organizations. *Administratie si Management Public*, (31), 6-20, DOI:10.24818/amp/2018.31-01

Jolliffe, I.T. (2002). 2nd Edition. *Principal Component Analysis*. New York: Springer-Verlag.

Karlis, D., Saporta, A. & Spinakis, A. (2003). Simple rule for the selection of principal components. *Journal Communications in Statistics - Theory and methods 32*(3), 643-666.

Ključnikov, A., Mura, L. & Sklenár, D. (2019). Information security management in SMEs: factors of success. *Entrepreneurship and Sustainability Issues 6* (4), 2081-2094.

Kral, P., Janoskova, K. & Kliestik, T. (2018). Key determinants of the public transport user's satisfaction. *Administratie si Management Public*, (31), 36-51, DOI: 10.24818/amp/2018.31-03

Lazaroiu, G., Kovacova, M., Kliestikova, J., Kubala, P., Valaskova, K., Dengov, V.V. (2018). Data governance and automated individual decision-making in the Digital Privacy General *Data Protection* Regulation. *Administratie si Management Public*, (31), 132-142, DOI: 10.24818/amp/2018.31-09

Mura**,** L., Vlacseková**,** D**.** (2018). Motivation of public employees: case study of Slovak teaching and professional staff. *Administratie si Management Public*, (31), 67-80, DOI: 10.24818/amp/2018.31-05

Mura, L., Daňová, M., Vavrek, R., Dúbravská, M. (2017). Economic freedom – classification of its level and impact on the economic security. *AD ALTA-Journal of Interdisciplinary Research, 7* (2), 154 – 157.

Novak, J., Purta M., Marciniak, M., Ignatowicz, K., Rozenbaum, K. & Yearwood, K. (2018). *The rise of Digital Challengers – How digitization can become the next growth engine for Central and Eastern Europe*. Digital McKinsey. McKinsey & Company.

Ragnedda, M., Kreitem, H. (2018). The three levels of digital divide in East EU countries. World of Media. *Journal of Russian Media and Journalism Studies*, *1* (4), 5-26.

Rakotomalala, R (2013). *Tanagra tutorial for PCA.* Tanagra Tutorials. URL:http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/en_Tanagra_Nb_Components_PCA.pdf

Saporta, G. (2006). *Probabilités, analyses des données et Statistiques*. Dunod.

Sebestova, J., Majerova, I. & Szarowska, I. (2018). Indicators for assessing the financial condition and municipality management. *Administratie si Management Public,* (31), 97-110, DOI: 10.24818/amp/2018.31-07

Shkarlet, S., Dubyna, M., Hrubliak, O., Zhavoronok, A. (2019). Theoretical and applied provisions of the research of the state budget deficit in the countries of Central and Eastern Europe. *Administratie si Management Public*, (32), 120-138. DOI: 10.24818/amp/2019.32-09.

Shpak, N., Odrekhivskyi, M., Doroshkevych, K., Sroka, W. (2019). Simulation of innovative systems under Industry 4.0 conditions. *Social Sciences*, 8, 202; doi:10.3390/socsci8070202

Ślusarczyk, B.(2018). Industry 4.0 – Are we ready? *Polish Journal of Management Studies*, *17*(1), 232-248. DOI: 10.17512/pjms.2018.17.1.19

Stverkova, H., Pohludka, M., Kurowska-Pysz, J., Szczepańska-Woszczyna, K.(2018). Cross-border enterprepreneurship in euroregion beskydy. *Polish Journal of Management Studies, 18*(2), 324-337. DOI: 10.17512/pjms.2018.18.2.26

Tvaronavičienė, M. (2019). Insights into global trends of capital flows' peculiarities: emerging leadership of China. *Administratie si Management Public*, (32), 6-17. DOI: 10.24818/amp/2019.32-01.

Ungerman, O., Dedkova, J. & Gurinova, K. (2018). The impact of marketing innovation on the competitiveness of enterprises in the context of Industry 4.0. *Journal of Competitiveness*, *10* (2), 132-148. http://doi.org/10.7441/joc.2018.02.09

Vezzani, A., Pugliese, E. & Gkotsis, P. (2018). *EU regions and the upgrading for the digital age*.
Luxembourg: Publications Office of the European Union.

Yeomans, K.A., Golder, P.A. (1982). The Guttman-Kaiser Criterion as a Predictor of the Number of Common Factors. *Journal of the Royal Statistical Society. Series D (The Statistician)*. *31* (3), 221-229.